

Authority linking with identifiers

Lessons learned on a major configuration change

IGELU Conference, September 11th 2023

Michael Hertig

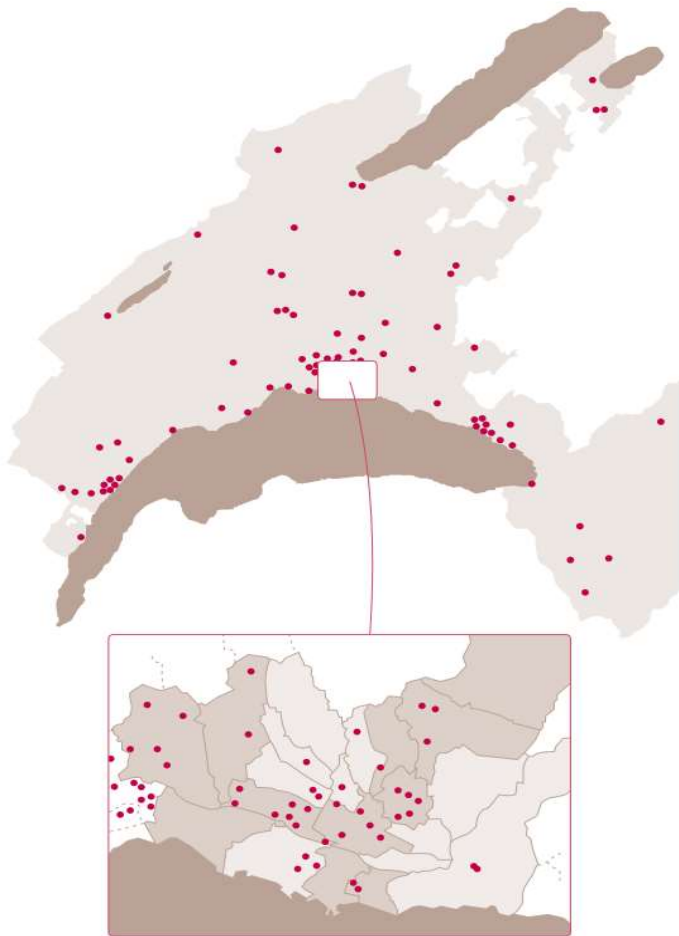


The Renouvaud Library Network

- Vaud Canton, Switzerland
- Capital city: Lausanne
- French speaking
- Bibliothèque Cantonale Universitaire Lausanne



The Renouvaud Library Network



- 149 libraries
 - Primary and high school, higher education, research, public, heritage, etc.
- ~ 200'000 patrons
- Union catalogue
 - 4.2M physical resources
 - 1.4M electronic resources

Authority vocabularies in Renouvaud

- Authority vocabularies currently in use

| Vocabulary | Function | MARC fields | Management in Alma | Coordination |
|----------------|------------------|--------------------|--------------------|------------------|
| ATC | Names | 1XX, 7XX, 8XX | NZ | NA |
| Rnv-mat | Names | 6XX | NZ | Pre-coordinated |
| RAMEAU | Subjects | 6XX | NZ/CZ | Pre-coordinated |
| MeSH | Subjects | 6XX | CZ | Post-coordinated |
| Jurivoc | Subjects | 6XX | CZ | Post-coordinated |
| IdRef | Names + Subjects | 1XX, 6XX, 7XX, 8XX | CZ | Pre-coordinated |

Authorities and bibliographic headings

- Authority control
 - Bibliographic heading content refers to authority records
 - If authorized form changes → bibliographic heading content changes
 - **Links**
- Enrichment in published data
 - Primo
 - Publishing profiles

```
100 0 $$a Costa-Gavras $$d 1933-....
375 $$a masculin $$2 unimarc
377 $$a fre $$2 639-2
400 1 $$a Gavrás, Konstantínos $$d 1933-.... $$w r $$4 f|xxx $$i Nom à l'état civil
400 1 $$a Gavrás, Kóstas $$d 1933-.... $$w r $$4 f|xxx $$i Nom à l'état civil
400 1 $$a Gavras, Constantin $$d 1933-....
400 1 $$a Γαβράς, Κώστας $$d 1933-....
```

```
100 1 $$a Costa-Gavras $$d 1933-.... $$0 (IDREF)067228801 $$4 cre
245 1 0 $$a Etat de siège / $$c Costa-Gravas, Franco Solinas
264 1 $$a Paris : $$b Stock, $$c 1973
```

1-1 de 1 Résultats ▾

1  LIVRE **Etat de siège**    

Costa-Gavras, Constantin Solinas, Franco
Paris : Stock 1973

 Disponible >

Alma default authority control

- Linking strategy = 'text'
- Alma linking job
- Alma Preferred Term Correction job (PTC)
- Authority Control Task List (ACTL)

Facets <<

Vocabulary >

Field v

648 (1)
100 (4)
830 (38)
810 (24)
[+ More \(15\)](#)

Mode v

Undefined (8108)

Tag v

Undefined (8108)

Authority Control Task List (1 - 20 of 8,108)

| | <input type="checkbox"/> | Report Type | Change Date | MMS ID | Title |
|----|--------------------------|--|-------------|---------------|------------------------------|
| 16 | <input type="checkbox"/> | Linking - BIB heading found no matching AUT headings | 03.09.... | 9910230648... | Études rurales. |
| 17 | <input type="checkbox"/> | Linking - BIB heading found no matching AUT headings | 01.09.... | 9910241345... | Étude en noir : tr de M. X / |
| 18 | <input type="checkbox"/> | Linking - BIB heading found no matching AUT headings | 02.09.... | 9910235664... | Étude des styles mobilier / |
| 19 | <input type="checkbox"/> | Linking - BIB heading found no matching AUT headings | 05.09.... | 9910241844... | Étranges dispari |
| 20 | <input type="checkbox"/> | Linking - BIB heading found no matching AUT headings | 06.09.... | 9910241838... | Éloi joue dans le / |

Changing authority control configuration



Why on earth would you do that?

- Character strings → Identifiers
- Transition to IdRef
 - Updating Renouvaud bib headings to IdRef bib headings
 - Renouvaud bib headings will better match Renouvaud authorities
- Better structured data
- Id's are the way to go!

A diagram illustrating the transition from character strings to identifiers. A light blue box labeled "URI" has a red arrow pointing to a red-bordered box containing the URL "http://catalogue.bnf.fr/ark:/12148/cb122733850". This URL is highlighted in a MARC record snippet. Below it, another red-bordered box contains the identifier "(IDREF)031537944". The MARC record snippet is as follows:

```
024 7 $$0 http://catalogue.bnf.fr/ark:/12148/cb122733850 $$2 bnf
035 $$a (IDREF)031537944
100 1 $$a Giacovelli, Enrico $$d 1958-....
```

Changing linking strategy

1. Change linking strategy parameter in Alma
 - 'text' → 'id'

2. Insert identifiers
 - Development of *Correctref*



Inserting identifiers in bib headings

- Use of \$0 subfield with authority ID
- Subfields \$0 AFTER the corresponding subfield(s)
- Some headings refer to more than one authority
 - Pre-coordinated subject headings
- Some headings also have non-controlled subfields
 - Name-title, congresses, etc.

700 1_ \$a Wagner, Richard, \$d 1813-1883

→

700 1_ \$a Wagner, Richard, \$d 1813-1883 \$0 (RNV_A)0000757412

711 _2 \$\$a Convegno liturgico internazionale \$\$n (17 : \$\$d 2019 : \$\$c Bose)

→

711 _2 \$\$a Convegno liturgico internazionale \$\$0 (RNV_A)0001624228 \$\$n (17 : \$\$d 2019 : \$\$c Bose)

Inserting subfields \$0 in bib headings

- ID insertion process from 5th to 16th February 2022
- ATC, IdRef, Rameau, MeSH, Jurivoc
- LCSH, LCNAF, FAST, GND
- 2.6M records (non CZ records)

Inserting subfields \$0 in bib headings

- Name headings (February 2022)

| Field tag | ATC | % |
|--------------|----------------|--------------|
| 100 | 1935051 | 97.70 |
| 110 | 26 | 51.43 |
| 111 | 1 | 16.67 |
| 130 | 284 | 9.49 |
| 700 | 1675865 | 97.35 |
| 710 | 276681 | 95.90 |
| 711 | 23657 | 94.32 |
| 730 | 461 | NA |
| 800 | 112574 | 0.29 |
| 830 | 672613 | 82.24 |
| Total | 4871584 | 92.52 |

Inserting subfields \$0 in bib headings

- Subject headings (February 2022)

| Field tag | Total | % |
|------------|-----------|-------|
| 600 | 92'598 | 60.23 |
| 610 | 26'184 | 49.41 |
| 611 | 1'205 | 61.08 |
| 630 | 21'587 | 34.97 |
| 648 | 4'715 | 9.08 |
| 650 | 1'156'693 | 87.66 |
| 651 | 161'205 | 87.04 |
| 655 | 185'466 | 91.08 |

| Vocabulary | Total | % |
|-------------------|---------|--------------|
| Fast | 74'129 | 94.41 |
| GND | 23'940 | 83.73 |
| Jurivoc | 439'206 | 94.98 |
| LCNAF/LCSH | 445'695 | 59.62 |
| MeSH | 167'605 | 96.43 |
| rnv-ram | 499'438 | 93.36 |


Changing Alma configuration

- Parameter 'authority_linking_strategy' = 'text' → 'id'
- Change done on February 17th
- Testing in Sandbox difficult
 - Authority jobs must be run manually

Changing linking strategy - Consequences



Discovery

- Primo - Prima! 

[< RETOUR](#)

ENVOYER À

DÉTAILS

OBTENIR D'UNE ... [Détails](#)

OBTENIR

LIENS

| | |
|---------------------------|---|
| Titre | Etat de siège |
| Par | Costa-Gravas, Franco Solinas |
| Auteur·e | Costa-Gavras, Constantin > |
| Contributeur·trice | Solinas, Franco > |
| Maison d'édition | Paris : Stock |
| Date | Publication : 1973 |
| Format | 296 p. : ill. ; 21 cm |
| Collection | Lutter ; 2 > |
| Note | Bibliogr.: p. 287-292 |
| Langue | Français |
| Source | Renouvaud |

[Liens](#)

[Informations bibliographiques](#)  [>](#)

Link on save

- Need to re-index the data
 1. Disable the authority jobs
 2. Full Inventory Indexing Job
 - a) Bibliographic data
 - b) Authority data
 3. Unlink bib records from authority records (43 Hours 22 Minutes 29 Seconds !)
 4. Reactivate the authority jobs
- Link on save works, but...

Preferred Term Correction

- Unwanted corrections !
- Alma replaces the whole heading by the authorized form:

700 12 \$\$a Pullman, Philip \$\$0 (RNV_A)0000650914 \$\$t **A la croisée des mondes**

→ \$\$a Pullman, Philip \$\$0 (RNV_A)0000650914

711 _2 \$\$a Convegno liturgico internazionale \$\$0 (RNV_A)0001624228 \$\$n **(17 : \$\$d 2019 : \$\$c Bose)**

→ \$\$a Convegno liturgico internazionale \$\$0 (RNV_A)0001624228

- About 2 weeks to restore the data !

Headings related to CZ vocabularies

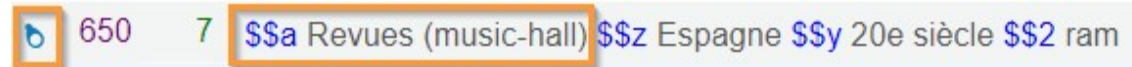
- For some CZ vocabularies subject headings did not link
 - MeSH, GND, Jurivoc, FAST
- Need for a space between the prefix and the ID :
 - 650 _7 \$\$a Cancer \$\$x Treatment \$\$0 (OCoLC) **fst00845538** \$\$2 fast
- Inconsistency
 - Link on save : need for space
 - F3 : no space
- Ex Libris solved the inconsistency efficiently!



Summing up: Alma and ID linking

- Constructed headings not supported
- Name-title headings not supported

- No partial linking

A screenshot of a library record snippet. The text is: `650 7 $$a Revues (music-hall) $$z Espagne $$y 20e siècle $$2 ram`. The field `650` is highlighted with a blue box, the field `7` is highlighted with a green box, and the field `$$a Revues (music-hall)` is highlighted with an orange box.

- No ID - no link
 - Imported bib records?

→ Need for adjustments

Introducing Correctref

- Correctref as third party authority control application
 - ATC, rnv-mat, Rameau (August 2022)
 - IdRef (October 2022)
- Disable the PTC job
 - ATC, rnv-mat, Rameau, LCNAF, LCSH, IdRef
 - Authority control rules
- Keep the Linking job
- Alma:
 - MeSH, Jurivoc



Correctref



Correctref sequential analysis

- Selecting the “best” matching authority heading



Corrections of bib headings

| | |
|-----------------|---|
| id-insertion | Insertion of a \$0 with ID in the bib heading |
| id-moving | Moving the \$0 with ID right after the linked subfield sequence |
| punct-update | Change of punctuation |
| alt-form-update | Updating the bib heading form according to the linked authority |
| code update | Updating the subfield codes according to the linked authority |

Corrections of bib headings

Preferred term correction and punctuation update on a name-title heading

700 12 \$a Fabricius, Fritz. - \$0
(RNV_A)0000117521 \$t <<Die
>>Rechtsstellung des Arbeitnehmers im
Staat der Industriegesellschaft

700 12 \$a Fabricius, Fritz, \$d 1919-2006. - \$0
(RNV_A)0000117521 \$t <<Die
>>Rechtsstellung des Arbeitnehmers im Staat
der Industriegesellschaft

Preferred term correction in constructed subject headings

650 _7 \$a Danse \$x Aspect religieux \$y
Moyen âge \$v Thèses et écrits
académiques. \$2 ram

650 _7 \$a Danse (religion) \$0
(RNV_A)1011941381 \$y Moyen âge \$0
(RNV_A)1011975717 \$v Thèses et écrits
académiques \$0 (RNV_A)1011933610 \$2 ram

Correctref reports

- Backoffice application
 - No Frontend

3 cases when Correctref cannot link:

1. Ambiguous heading: the heading matches more than one authority
 2. Missing authority record: Correctref could not match the bib heading with any existing authority
 3. Inconsistent ID: the authority ID and the form in the bib heading do not match
- Reports on demand (Excel files)

Conclusions



Conclusion

- Changing linking strategy
 - Fundamental change in metadata management
 - Difficult technical part
 - Difficult testing
 - Heavy consequences
- Need for support
 - Checklist?
- Need for efficient authority control
 - Third party application
 - Still relying on the Alma linking job
 - Partial enrichment of headings in Primo and publishing profiles

Renouvaud

Réseau vaudois
des bibliothèques

**Thanks for your
attention!**

Michael.hertig@bcu.unil.ch



Agenda

- Context
- Changing authority control configuration
- Consequences
- Correctref
- Conclusions

Authorities and bibliographic headings

- Bibliographic headings → Load-bearing structure
- Authorities → Foundation
- **Don't mess with it!**



Correctref standing job

- Correctref authority control (August 2023) – Name headings
- 2.6M records

| Field tag | idref | Idref % | ATC | ATC % | Total |
|-----------|-------|-------------|---------|--------------|-------|
| 100 | 37374 | 1,8% | 1966461 | 95,8% | 97,6% |
| 110 | 3 | 5,7% | 26 | 49,1% | 54,7% |
| 111 | 1 | 7,7% | 0 | 0,0% | 7,7% |
| 130 | 0 | 0,0% | 291 | 9,2% | 9,2% |
| 700 | 39648 | 2,2% | 1715688 | 93,9% | 96,1% |
| 710 | 2796 | 0,9% | 298822 | 92,3% | 93,2% |
| 711 | 216 | 0,8% | 23705 | 93,2% | 94,1% |
| 730 | 0 | 0,0% | 484 | 1,6% | 1,6% |
| 800 | 1097 | 0,9% | 112846 | 95,1% | 96,0% |
| 830 | 0 | 0,0% | 680573 | 83,2% | 83,2% |
| Total | 81135 | 1,3% | 4798896 | 79,5% | 80,8% |

Correctref standing job

- Correctref authority control (August 2023) – Subject headings

| Field tag | Rnv-mat/ram | % |
|-----------|-------------|--------------|
| 600 | 46215 | 81,4% |
| 610 | 5699 | 83,4% |
| 611 | 721 | 96,5% |
| 630 | 6113 | 83,8% |
| 648 | 35 | 97,2% |
| 650 | 558712 | 95,8% |
| 651 | 28018 | 88,3% |
| 655 | 147253 | 95,8% |
| Total | 792766 | 90,3% |

Changing linking strategy

- 'text' → 'id'
- Insertion of identifiers
 - Subfields \$0
 - Development of *Correctref*
- ID insertion (February 2022):
 - ATC, IdRef, Rameau, LCSH, LCNAF, FAST, GND, MeSH, Jurivoc

